스케일러블 비디오 부호화 표준화 동향

최해철 | 강정원 | 배성준 | 유정주
한국전자통신연구원

요 약

근래의 무선망 및 인터넷의 초고속화에 따라 다양한 멀티미디어 서비스가 활성화 되고 있으며, 특히 BcN(Broadband convergence Network) 및 통방융합망의 등장으로 압축 부호화 기술만을 개발하던 시기와 달리 멀티미디어의 생성, 전송 및 소비 환경의 다양한 조건들에서 QoS를 보장하기 위해 비디오 부호화의 스케일러빌러티를 제공하기 위한 표준화가 진행되고 있다. 본 고에서는 스케일러블 비디오 부호화의 개요에 대해 설명하고, 현재 진행 중인 SVC(Scalable Video Coding)의 표준화 동향을 기술한다. 그리고 SVC에 표준으로 채택된 기술 중 시간/공간/화질적 스케일러빌러티를 제공하기 위한 대표적인 알고리즘에 대해 설명한다.

I. 서 론

근래의 무선망 및 인터넷의 초고속화에 따른 정지영상 및 동영상 등의 정보를 포함한 다양한 멀티미디어 서비스가 통방융합 현상과 더불어 새로운 패러다임으로 활성화 되고 있다. 우리는 이러한 광대역 시대를 맞아 통신회선이 고속화 되는데, '압축기술이 왜 필요한가? 압축기술의 배경이 무엇인가?'에 대한 의문을 자연스레 가지게 된다. 이것은 비디오 코덱 기술이 발전해온 지난 20년간 반복되어 나오는 질문이다. 확실히 초고속 통신망 시대를 맞아서 유선 또는 무선의 광대역 접속회선을 싸게 사용할 수 있는 환경이 이루어 졌다. 댁내의 접속회선은 FTTH (Fiber-To-The-Home)으로 초고속화 되고, 저장매체도 DVD 레코더와 같은 대용량 제품이 등장했지만, 우리들이 사용할 수 있는 문자, 음성, 사진, 동영상과 같은 멀티미디어 정보량 역시 보다 좋은 색상, 화질이 요구되어, 필요한 정보량 역시 시대에 따라 증가해 가고 있다. 더욱이 아프리카(Afreeca), U-Tube 등 근래에 유행하는 UCC(User Created Content)는 소비자의 이러한 요구가 증폭되고 있음을 대변한다. 또한 BcN(Broadband convergence Network) 및 통방융합망의 등장으로 압축 부호화 기술만을 개발하던 시기와 달리 최근의 디지털 멀티미디어의 생성, 전송 및 소비 환경은 광범위하게 급변하고 있다. 예를 들어 서로 다른 특성 전송 용량 및 에러 발생율을 갖는 이종망이 혼재해 있으며, DMB(Digital Multimedia Broadcasting) 서비스와 같은 통신과 방송의 융합화가 빠르게 진행되고 있다. 더욱이 비디오를 소비하는 단말 또한 다양한 화면 크기와 연산 능력을 가진 새로운 제품들이 시장에 쏟아져 나오고 있다. 현재의 비디오 압축 기술은 주로 고정된 영상 포멧과 비트율(bit-rate)로 압축/전송하는 단일 계층 압축 부호화(single layer video coding) 방식이다. 즉 주변에서 흔히 볼 수 있는 mp4, avi, wmv 형식의 비디오 파일들은 고정된 영상 크기 및 화질을 가지고 있다. 앞서 언급한 환경에서 이러한 단일 계층 압축 부호화 방식은 멀티미디어의 생성에서부터 사용자가 직접 시청하는 단말의 소비까지 끊김 없는 품질의 보장에 대한 사용자의 요구를 효과적으로 충족시킬 수 없다. 이러한 환경에서 각 네트워크의 조건과 비디오를 시청하는 단말기의 성능에 맞추기 위해 변환부호화(transcoding)[1]나 비트량변환(transrating) 방식등이 개발

되어 왔다. 그러나 이 방식들은 제한된 범위에서의 영상 형식 및 비트율 적응이 가능하지만 큰 복잡도로 인해 실시간 적응이 요구되는 환경에서 널리 쓰이지 못했다. 따라서 언제 어디서나 어떠한 단말로든 효과적으로 동영상을 전송/소비하기 위해서는 앞서 기술한 고압축 특성을 가진 부호화 기술뿐만 아니라 전송 및 소비 환경에 따라 적응적으로 변환할 수 있는 스케일러빌러티(scalability)라는 개념이 생겨났으며 최근 이에 대해 활발한 연구가 진행 중이다. 동영상에 있어서 스케일러빌러티 기능을 위한 기존 연구는 MPEG-2 [2] 와 MPEG-4 (part2) [3][4] 비디오에서 계층 부호화에 기반한 시도가 있었으나 하나의 계층으로 부호화하는 기존 압축 부호화 기법에 비해 부호화 효율이 낮고, 여러 종류의 스케일러빌러티를 종합적으로 지원하지 못하는 취약점이 있었다. 이를 해결하면서 실시간 비디오 전송에서의 다양한 스케일러빌러티를 지원하기 위해 현재 ISO/IEC Moving Picture Experts Group (MPEG)과 ITU-T Video Coding Experts Group (VCEG)의 Joint Video Team (JVT)은 Scalable Video Coding (SVC) [5][7][7] 표준화를 MPEG-4 part 10 AVC/H.264 [8]의 수정판(Amendment)으로 진행하고 있으며, 여기에는 유럽, 미국, 한국이 적극적으로 참여하여 기여하고 있다. 본 고에서는 스케일러블 비디오 부호화의 개요에 대해 설명하고, 현재 표준화가 진행 중인 SVC의 표준화 동향을 기술한다. 그리고 SVC에 표준으로 채택된 기술 중 시간/공간/화질적 스케일러빌러티를 제공하기 위한 대표적인 알고리즘에 대해 설명한다.

II. 스케일러블 비디오 부호화 개요 및 표준화 동향

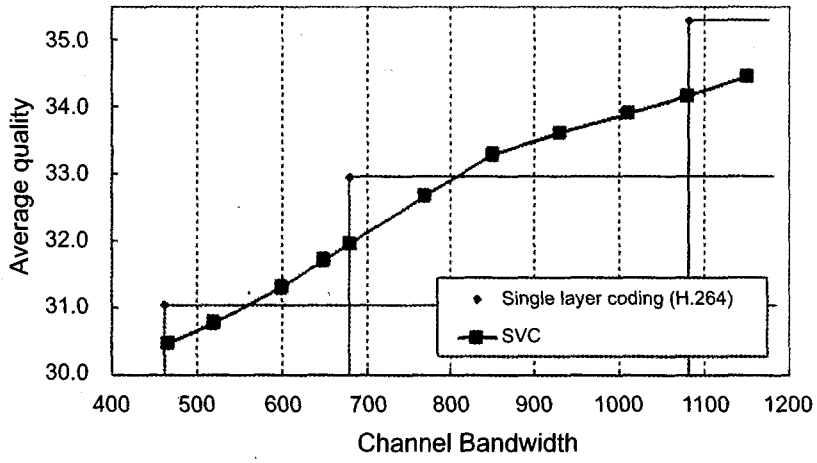
1. 스케일러블 비디오 부호화 개요

네트워크를 통한 멀티미디어 서비스에서 스케일러블 비디오 부호화의 목적은 특정 비트율로 화질을 우화하는 것이다. 이 때 그 비트스트림(bit-stream)은 임의 비트율에서도 복호화가 되어야 한다. 단일 계층 압축 부호화기는 하나의 비트율/프레임율/영상크기만을 지원하는 하나의 비트스트림만을 생성하는데 반해, 스케일러블 비디오 부호화기는 다양한 비트율/프레임율/영상크기에 대한 스케일러빌러티를 지원 한다. SVC 표준에서 스케일러빌러티(scalability)란 'H.264/MPEG4-AVC [7] 부호화와 비교 시 비슷한 비트율-왜곡(Rate-Distortion) 성능을 가짐과 동시에 공간, 시간, 화질 측면에서 비트스트림의 한 부분을 제거할 수 있는 기능(functionality)'으로 정의한다. SVC는 여러 개의 비디오 계층을 하나의 비트열로 부호화하며, 각 층은 각각의 비트율, 프레임율, 영상 크기 및 화질을 가지고 있다. 즉 비디오를 공간적/시간적/화질적 차원의 임의 값을 가지는 비트열로 부호화하며, 그 세 가지 차원의 조합에 따라 폭넓은 스케일러빌러티를 제공할 수 있다. 기술적인 관점에서, 하나의 스케일러블 비트스트림은 두 개 혹은 그 이상의 의존적인 계층으로 구성될 수 있다. 이 경우, 스케일러블 코덱(codec)은 하나의 기본 계층과 스케일러블 상위 계층들로 구성된다. 여기서 기본 계층 및 연속되는 상위 계층의 정보가 함께 이용되어 보다 개선된 비디오 비트스트림을 만든다. 그 예로 화질 스케일러빌러티는 하나의 비트열로부터 동일한 공간 및 시간 축(dimension)을 갖지만 각각 다른 화질을 갖는 비디오 계층들을 만들어 낸다. 일반적으로 기본 계층은 기본적인 비디오 화질을 제공하고, 연속된 상위 계층은 이전 계층들로 만들어진 비디오보다 높은 화질을 갖도록 부호화한다. 마찬가지로 시간 및 공간 해상도에서도 동일한 원리를 적용하여 스케일러빌러티를 지원한다. 하나의 차원에 대한 스케일러빌러티로는 최적의 가변성과 적응성을 제공할 수 없는 응용분야가 있을 수 있으며, 이는 시간, 공간, 화질 측면 모두의 스케일러빌러티를 필요로 한다.

(그림 1)에서의 비트율-왜곡 곡선(distortion-rate curve)은 MPEG-4 AVC/H.264와 SVC 기술을 이용할 때의 비트율 대비 화질 곡선을 보여준다. 그림에서 단일 계층 압축 부호화 방식은 채널의 대역폭이 증가했음에도 불구하고 더 좋은 화질을 제공할 수 없지만, SVC는 대역폭에 따라 보다 좋은 화질의 비디오를 지원할 수 있다. SVC 의 목표는 단일 계층 부호화 방식의 R-D 곡선과 일치하는 연속적인 곡선을 얻는 것이다. 실제 기존 H.264와 비교 시 10% 내외의 비트율을 허용한다.

2. 스케일러블 비디오 부호화 표준화 동향

공간/시간/화질적 스케일러빌러티는 가장 기본적인 스케



(그림 1) 비디오 부호화 방식에 따른 채널 대역폭별 화질 성능 그래프

일러블 기능으로서 MPEG-2와 MPEG-4에 모두 포함되어 있다. 근래의 스케일러빌러티 기능 연구로 MPEG-4 FGS (Fine Granularity Scalability)가 있다. MPEG-4 FGS[5]는 세밀한 간격으로 비트율을 조절할 수 있을 뿐만 아니라 전송하고자 하는 비트율은 압축하는 시점이 아닌 전송하는 시점에 결정할 수 있게 된다. 따라서 여러 비트율로 컨텐츠를 전송하는 경우에 하나의 압축된 비트스트림으로부터 여러 비트율의 전송용 비트스트림을 매우 손쉽게 생성하는 것이 가능해졌다. MPEG-4 FGS 표준화가 종료된 이후, MPEG에서는 MPEG-21 (part 13)을 시작점으로 SVC라는 소그룹(AdHoc)을 결성하였다. 2004년의 CFP(Call for Proposal) [5]에서 다음 두 가지 범위의 시나리오를 설정하고 제안 기술을 검증하였다. 첫 번째 시나리오는 넓은 범위의 스케일러빌러티로 감시시스템, 방송, 저장장치를 적용대상으로 삼고 있으며, 1.2초 범위의 임의 접근(random access)을 요구한다. 두 번째 시나리오는 보다 제한된 범위의 스케일러빌러티로 스트리밍(streaming), 이동통신 등의 분야에 관한 것이며 임의 접근의 제한은 없다.

기본적인 SVC의 프레임워크로 DCT 기반의 방식과 웨이블릿(wavelet) 기반의 방식이 제안 검증 되었다. 검증 결과, 웨이블릿 기반 기법이 근본적으로 스케일러빌러티를 포함하고는 있으나, 웨이블릿 필터의 제한으로 인해 압축 효율성 측면에서 독일 HHI에서 제안한 DCT 기반의 기법보다 낮은 성능을 보여주었다. 따라서 DCT 기반 기법이 SVC의 기본 프레임워크로 채택 되었으며, 이를 기반으로 많은 알고리즘들이 연구, 개발되고 있다.

지난 2005년 1월 MPEG 홍콩 회의에서, 또 다른 비디오 국제 표준화 단체인 ITU-T의 VCEG이 SVC 표준화에 동참하면서 JVT에서 스케일러블 비디오 표준화가 수행 중이다. 2006년 1월에 Committee Draft가 발표되어 있으며, 2007년 4월에는 FGS를 제외한 시간적/공간적 화질에 대한 프로파일에 대해 FDAM(Final Draft Amendment)가 예정되어 있고, 2008년 4월에 FGS를 포함한 프로파일에 대해 FDAM을 발표할 예정이다. 2006년 4월 현재 JD(Joint Draft) 9 [6]과 JSVM(Joint Scalable Video Model) 9 [7]가 발표 되었다.

III. 스케일러블 비디오 부호화 기술

SVC의 부호기에 대해 기술하고 있는 JSVM-9 [7]은 다음 두 가지 방법을 통해 스케일러블 기능을 제공한다. 한 가지는 비트플레인(bit-plane) 기반 산술코딩(arithmetic coding)과 같은 부호화 방법이고, 다른 하나는 기존의 여러 스케일러블 표준에서 사용되었던 계층 기반 기법이다. JSVM에서는 시간적/공간적/화질적 스케일러빌러티를 모두 제공하기 위

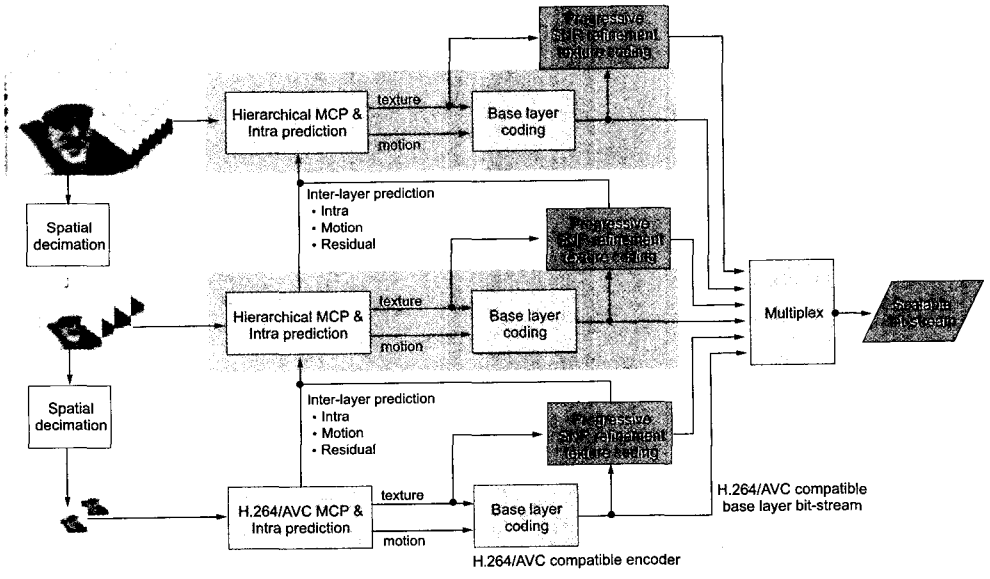
해 상기 두 가지 방법을 모두 사용하였다. 시간적 스케일러빌러티는 Hierarchical B picture [9] 구조로써 지원하고, 공간적인 스케일러빌러티 제공을 위해서는 down/up 표본화(sampling)를 통해 여러 해상도를 가지는 영상들을 구성한다. 그리고 화질(quality 혹은 SNR) 스케일러빌러티를 위해서는 CGS의 경우 공간 스케일러빌러티를 위한 기법과 동일한 방식에 양자화 계수(quantization parameter)만을 변경하며, FGS의 경우 2-scan 방식과 cyclic 부호화 기법을 이용하여 스케일러빌러티를 지원한다[10].

JSVM 부호기 구조는 제공하려는 스케일러빌러티의 종류 및 단계에 따라 달라진다. 3개의 공간적 계층으로 구성하는 경우 SVC 부호기 구조는 (그림 2)와 같다. 공간적 스케일러빌러티를 위해 입력 영상의 공간 해상도를 줄여 작은 크기의 영상을 하위 계층에 입력 시킨다. 이때 통상적으로 MPEG-4 AVC/H.264의 6-tap down-sampling 필터를 사용하지만, 이 필터에 대해서는 표준으로 규정하지 않고 있다. 그 후 공간적으로 분할된 각 계층의 입력 비디오 신호에 Hierarchical B picture를 적용하여 시간적 스케일러빌러티를 지원한다. 이 과정의 결과는 시간적인 저주파 영상 [L]과 고주파 영상 [H]의 집합, 잔여 텍스쳐 정보, 블록 단위의 움직임 정보이다. 움직임 정보와 잔여 텍스쳐 정보는MPEG-4 AVC/H.264 알고리즘을 기초로 부호화하면서, 부가적으로 화질 스케일러빌러티를 제공하기 위해 CGS(Coarse Grain Scalability), MGS(Medium Grain Scalability) 혹은 FGS(Fine Grain Scalability) 기법을 적용한다. 또한 전체적으로 각 공간 계층 간의 정보 중복성(redundancy)을 줄이기 위해, 움직임 정보를 그 자체로 부호화하지 않고 하위 계층에 있는 대응 블록의 움직임 정보와의 차이 혹은 그 수정 값을 포함시켜 부호화하는 방식을 이용할 수 있다. 잔여 텍스쳐(residual texture) 정보 및 화면내(intra) 매크로블록의 예측 과정에 대해서도 동일한 개념의 방식을 MPEG-4 AVC/H.264에서의 방식들과 함께 적용한다. 유의할 점은SVC는 MPEG-4 AVC/H.264의 수정본으로 표준화가 진행 중이기에, SVC의 공간적 기본 계층(그림에서 하단 블록, H.265/AVC compatible encoder)은 반드시 MPEG-4 AVC/H.264와 완벽한 호환성을 가져야 한다.

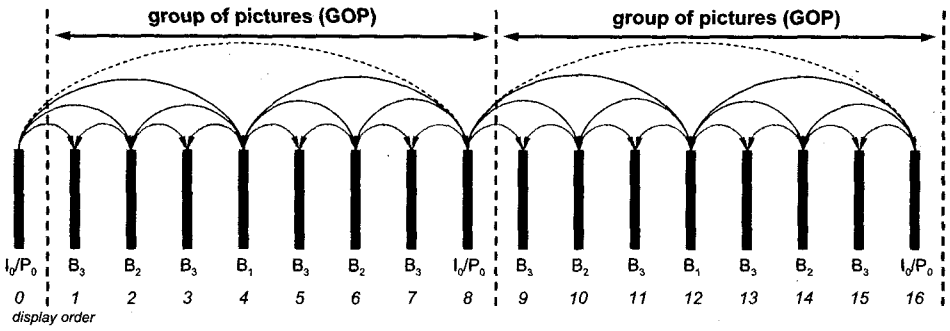
1. 시간 스케일러빌러티 관련 기술

1.1. Hierarchical B picture

AVC 에서 모든 영상은 참조 영상으로 사용할 수 있고, 다음에 오는 영상의 움직임 보상 예측에 이용할 수 있다.



(그림 2) 공간적 3 계층의 스케일러빌러티를 제공하는 스케일러블 부호기



(그림 3) GOP 크기가 8이며, 4개의 시간적 레벨을 갖는 hierarchical B picture 구조

Memory management control operation (MMCO) commands는 복호화된 영상을 저장할 수 있는 decoded picture buffer (DPB)를 적절하게 제어한다. 그리고 DPB에 저장된 다른 영상의 움직임 보상 예측에 이용되는 참조 영상들은 reference picture list re-ordering (RPLP) commands를 사용하여 임의적으로 선택이 가능하다. 이러한 AVC의 특징들은 이전 비디오 부호화 방식들이 지원하지 못했던 임의적인 부호화 및 참조 영상의 선택을 가능하게 한다. 이를 기반으로 B 영상의 참조를 허용하는 General B picture를 정의함으로써 시간적 스케일러빌러티를 지원하는 Hierarchical B picture가 제안되었다.

(그림 3)은 4 단계를 갖는 전형적인 hierarchical B picture 구조를 보여준다. 비디오의 첫 영상은 IDR (Instantaneous Decoder Refresh) 영상으로서 화면내 부호화된다. 이전에 부호화된 모든 영상들이 현재 부호화되는 영상보다 화면 표시 순서(display order) 상 앞서 위치할 때, 이것을 key picture(그림 3의 2번째, 3번째 계층의 블록)라고 부른다. 시간적으로 현재 key picture와 이전 key picture사이에 위치하고 있는 모든 non-key picture들을 하나의 group of pictures (GOP)라고 한다. Key picture는 임의 접근을 위해 화면내 부호화되거나, 이전 key-picture를 참조 영상으로 움직임 보상 예측을 통해 P 영상으로 화면간(inter) 부호화 된다. GOP의 나머지 영상들은 그림 3과 같이 계층적으로 참조된다.

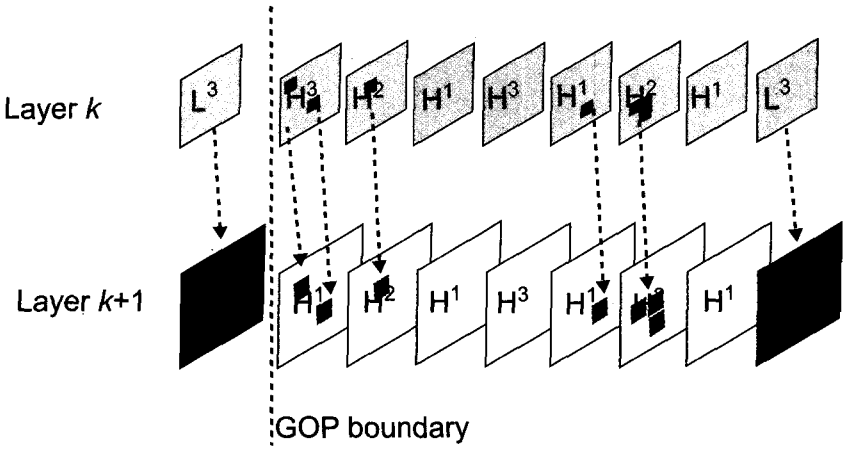
2. 공간적 스케일러빌러티 기술

공간적 스케일러빌러티는 각 영상 크기의 계층을 쌓는 피라미드 개념으로 지원한다. 이 때 각 해상도의 비디오는 하위 계층의 부호 결과물인 움직임, 텍스쳐 및 잔여 신호 정보를 이용 함으로써 새로운 계층으로 부호화된다. CGS, MGS 스케일러빌러티도 공간적 스케일러빌러티와 동일한 방식으로 지원한다. 단지 다른 점은 각 계층의 입력 영상 크기가 동일하는 것이다. 따라서 공간적 스케일러빌러티에서만 계층간 참조 시 움직임과 텍스쳐 정보가 이 영상 해상도의 비율에 따라 변화한다. 본 절에서는 각 공간적 또는 CGS 계층의 계층간(inter layer) 상관관계를 이용하기 위해서 제안된 다양한 계층간 예측 기술을 설명한다.

2.1. Inter-layer intra texture prediction

하위 계층의 정보를 사용하는 계층간 화면내 텍스쳐 예측은 Intra_Base 예측이라 하며 MB type이 Intra_BL로 설정된다. 현재 표준에서는 멀티 루프(multi-loop)는 지원하지 않고, 하위 계층의 대응 위치 블록이 화면내 부호화된 매크로블록 안에 위치할 때만 적용된다. 이 설정에서는 목표 영상 해상도를 위해 single decoding loop만이 쓰인다.

Intra_Base 예측에서 참조 신호를 얻기 위해서, 하위 계층의 동일 대응 지점에 위치한 블록의 화소값에 대해 de-blocking 필터링을 수행하고 매크로블록 단위의 경계 확장(border extension) 작업 후 (그림 4)와 같이 보간(interpolation)하게 된다. 보간 전 매크로블록 단위의 경계 확장(border extension) 작업을 하여 4-sample border로 확장 시킨다. 보간 과정은 상위 계층과 기본 계층간의 실제 공간 해상도 비율에 의해 결정된다. 이 비율이 2일 경우, 밝기 값에



(그림 4) 화면내 Macroblock을 위한 low-complexity 계층간 예측

대해서 보간 과정은 AVC의 half-pel 보간 필터를 사용되며 컬러 값에 대해서는 2-tap 필터를 이용한다. 반면에 이 비율이 1 또는 2가 아닌 non-dyadic 인 경우, 확장 공간 스케일러빌러티에서의 필터[6]가 이용된다. 결국 보간된 신호를 현재 블록의 참조 신호로 활용하여 그 차이 값만을 부호화 한다. 이 Intra_Base 예측은 SVC에서 부호화 효율을 높이는 주된 기술 중 하나이다.

2.2. Inter-layer motion prediction

대응 하위 계층이 화면간 부호화 되었다면 상위 계층의 블록은 하위 계층으로부터 움직임 정보 및 모드 정보를 참조할 수 있다. motion_prediction_flag_lx이 1인 경우, 상위 계층에서는 움직임 정보 및 매크로블록의 분할(partitioning) 정보를 별도로 부호화하지 않는다. 이 매크로블록 모드는 이전 계층의 대응 매크로블록의 분할정보 및 움직임/예측 정보가 그대로 이용되는 것을 의미한다. 이전 계층의 공간 해상도가 현재 계층보다 작을 때, 매크로블록 분할 정보와 움직임 벡터가 비율에 맞춰 조정된다. 현재 매크로블록의 분할을 위해, 이전 계층의 대응 위치 매크로블록/서브-매크로블록, 분할 지시자(partition index)와 같은 참조영상 지시자(reference index)가 사용된다. 이 때 움직임 벡터는 향상 계층과 기본 계층간의 해상도 비율을 곱하여 사용한다.

사용하지 않으면, 대응 참조영상 지시자, 움직임벡터, 매크로블록 모드는 AVC 규격(syntax)에 따라서 부호화된다.

2.3. Inter-layer residual prediction

한 계층에서 다음 계층으로 전달되는 움직임 정보가 변하면, 잔여 정보를 예측하는 것에 도움이 될 수도 있고 되지 않을 수도 있다. 만약 현재 계층 블록의 움직임 정보가 이전 계층의 대응 위치 블록의 움직임 벡터와 동일하거나 거의 유사하면, 이전 계층의 잔여신호가 현재 계층의 잔여 신호를 위한 예측으로 사용할 때 부호화 효율이 증대될 수 있다. 그래서 현재 잔여 신호와 이전 계층의 복호된 잔여 신호의 차가 부호화된다. 그러나 움직임 벡터가 유사하지 않거나, 매우 다를 경우에는 잔여 신호의 예측이 부호화 효율을 저하시킬 수 있다. 그러므로 잔여 신호의 예측을 다음의 flag로 선택하여 사용한다.

• Residual_prediction_flag

만약 이전 계층이 현재 계층의 절반인 공간 해상도를 가질 경우, 잔여 신호는 예측 신호로 사용되기 전에 separable bi-linear 필터를 사용하여 상향 표본화(up-sample)된다. 계층간 공간 해상도 비율이 1또는2가 아닐 경우, 보간 작업은 AVC의 quarter-pel을 기반으로 하여 작동된다. 보간 필터는 변환 블록 간의 경계를 건너서 적용되지 않는다.

3. 화질 스케일러빌러티 관련 기술

SVC는 화질적 스케일러빌러티를 위해서 그림 5와 같이 양자화에 의한 오차 신호를 이전 계층 보다 적은 양자화 계수

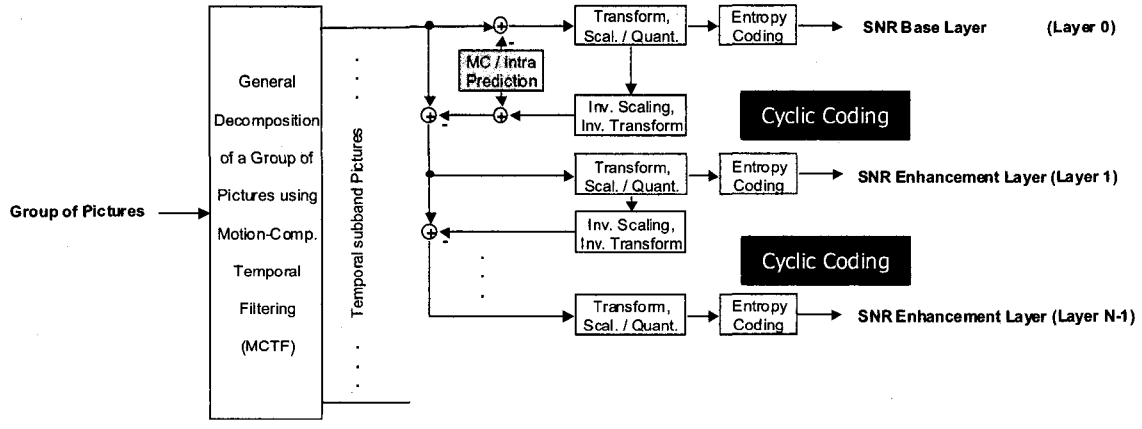
값으로 보정하여 부호화 하는 개념이다. 즉 첫번째 계층은 큰 양자화 간격으로써 듬성듬성하게 표현한 후, 계층이 올라갈수록 양자화 간격의 폭을 세밀하게 조정하여 보다 좋은 화질의 계층을 쌓는 개념이다. 기술적으로 살펴보면, 앞서 설명한 Hierarchical B picture 에 의한 시간적 분할 (decomposition) 이후의 신호들을 최고 양자화 계수값 (Quantization Parameter)으로써 기본 화질을 갖는 기본 계층을 부호화 한 후, 이 기본 계층을 다시 복호하여 원 신호와의 차이값을 구한다. 이 차이값, 즉 양자화 에러를 다시 DCT 변환하여 기본 계층보다 적은 양자화 계수(Quantization parameter)로 양자화 한다. 이때 QP는 이전 계층보다 6 적은 값을 선택하는데, 이는 H.264에서 QP 값이 6이 줄면 양자화 간격 크기가 반으로 감소되기 때문이다.(비트량은 약 두 배 증가) 이렇게 양자화 에러를 단계적으로 보정 부호화 하는 부호화 방식을 Progressive Refinement 라고 하며, 이렇게 부호화된 슬라이스를 Progressive Refinement (PR) slice라 한다. 이 PR 슬라이스는 2 scan 기법 및 cyclic 부호화 기법을 이용하여 부호화 함으로써 전체 프레임에서 각 블록별로 낮은 주파수의 값들을 순서대로 전송할 수 있다. 따라서 복호화 단에서 한 프레임에 속한 정보 중 후반부에 있는 데이터가 없더라도 복호화가 가능하다. 이 때문에 바이트 단위의 세밀한 화질 스케일러빌러티가 지원되는 것이다.

(그림 6)은 cyclic 부호화 방식의 예이다. 이 cyclic 부호화 방식의 목적은 한 블록 내의 정보를 전체 비트스트림으로 분산시켜 FGS 패킷의 임의 부분이 제거 되었더라도 전체 영상의 화질 차이가 적도록 하는 것이다. 이를 위해 사이클(cyclic)이라는 것을 정의하였다.

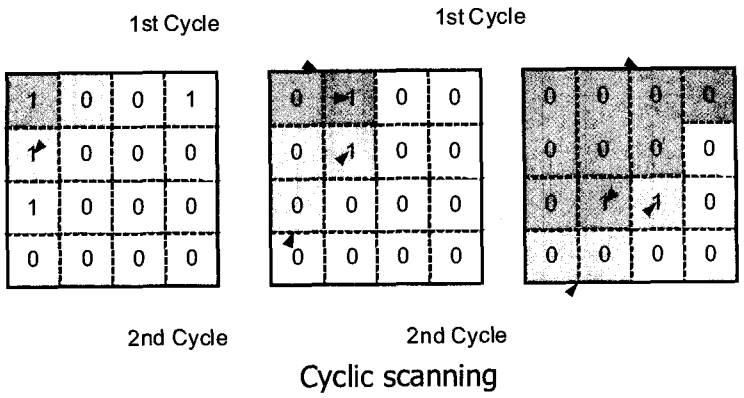
각 4x4 블록에서 1개의 0이 아닌 값 즉 significant 비트가 나오기 직전까지만 부호화 한 후, 다음 블록을 부호화한다. 이를 영상의 끝까지 진행하는 것을 하나의 사이클이라고 하며, (그림 6)에서는 2개의 사이클을 보여 준다. 우선 첫 번째 사이클을 살펴보면, 첫 번째 블록에서는 첫 번째 위치에 significant 값이 존재하므로 1개의 값만 부호화 되며, 2번째 블록에서는 2개의 값이 부호화 되고, 3번째 블록에서는 9개의 값이 부호화 된다. 이 사이클을 모든 블록 내의 모든 값이 부호화 될 때까지 반복하게 된다.

4. 복합 스케일러빌러티 (Combined Scalability)

SVC에서는 앞 절에서 설명한 공간/시간/화질 스케일러빌러티가 서로 유기적으로 사용될 수 있다. 예를 들면 비트스트림 CIF@30fps로부터 비트스트림 QCIF@30fps와 비트스트림 CIF@15fps를 추출할 수 있다. (그림 7)은 이러한 복합 스케일러빌러티의 한 예를 나타낸다. 이 그림에서 공간 해상도 기본 계층(QCIF)은 MPEG-4 AVC/H.264를 사용하여 15Hz로 부호화 되었다. 이때 시간적 스케일러빌러티를 제공하기 위해 Hierarchical B picture의 구조가 사용된다. 공간 상위 계층 (CIF)는4개의 시간적 분할 단계를 가지는 Hierarchical B picture를 사용하여 30Hz로 부호화되었다. 각



(그림 5) 화질 스케일러빌러티를 위한 부호화기 구조



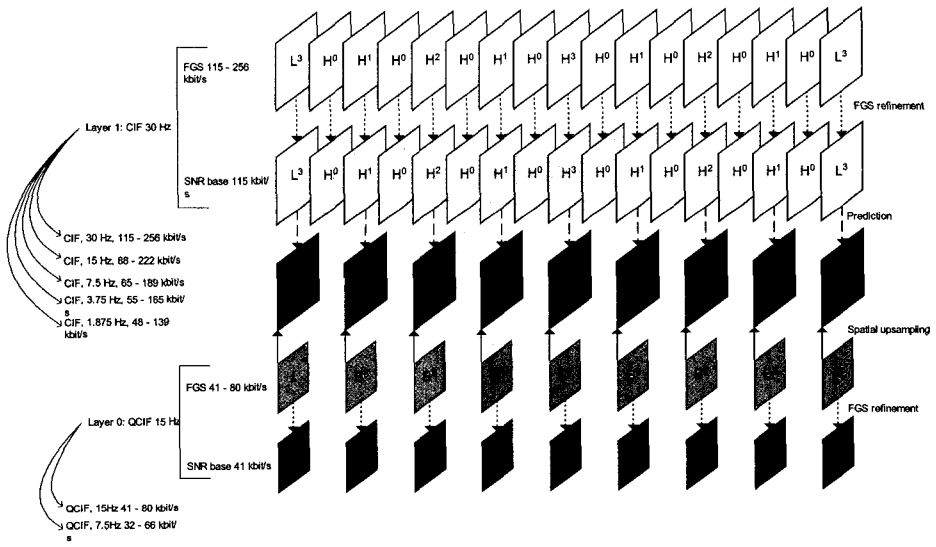
(그림 6) Cyclic 부호화 방식의 예

공간 계층은 FGS를 사용하여 SNR 스케일러빌러티를 제공한다.

비트율 측면을 보면, QCIF 영상 크기를 갖는 계층(layer 0)은 최대 80kbit/s로 부호화된다. SNR 스케일러빌러티를 이용하여 QCIF, 15Hz 비트스트림은 41kbit/s에서 80kbit/s 사이의 비트율로 전송 가능하며, 네트워크 상황이 악화되어 더 낮은 비트율 혹은 프레임율을 제공해야 한다면 B3 프레임들을 전송하지 않고 7.5 Hz의 비트스트림을 만들어 32kbit/s에서 66kbit/s사이의 비트율로 지원 가능하다. Layer 0를 포함하는 CIF 크기를 갖는 계층(layer 1)은 최대 256kbit/s로 부호화된다. 이 CIF, 30Hz 비트스트림은 115kbit/s에서 256kbit/s사이의 비트율로 다양하게 전송/소비될 수 있다.

시간적 스케일러빌러티를 위해 {H}0, {{H}0,{H}1}, {{H}0,{H}1,{H}2}, 와 {{H}0,{H}1,{H}2,{H}3}를 제거하여, 15, 7.5, 3.75와1.875Hz의 프레임율을 제공할 수도 있다.

서론에서 밝혔다시피 여러 가지 네트워크가 융합되며, 화면 크기 및 성능이 다양한 단말이 혼재해 있는 환경에서 위



(그림 7) 복합 스케일러빌러티의 예

와 같은 다양한 스케일러빌러티는 하나의 비트스트림으로 그 많은 종류의 조건들을 충족 시킬 수 있는 매우 유용한 기술임을 알 수 있다.

IV. 결 론

현재 ISO/IEC MPEG 과 ITU-T VCEG의 JVT에서는 멀티미디어를 다양한 망 및 단말 환경에서 서비스할 때 QoS를 보장하기 위한 목적으로 스케일러빌러티를 제공하기 위한 Scalable Video Coding에 대한 표준화를 진행하고 있다. 따라서 본 고에서는 스케일러빌러티의 필요성과 SVC의 개요에 대해 소개하고, 이의 표준화 동향과 스케일러빌러티를 지원하기 위한 대표적인 기술들에 대해 설명하였다.

결론적으로 SVC는 이종망 환경에서 발생하는 사용자의 선호도, 소비 단말기와 네트워크의 조건 등 다양한 전달 및 소비 환경에 효과적으로 대처할 수 있으며, 스케일러빌러티를 갖는 계층들의 조합을 하나의 비트열로 부호화하여 적응적인 전달/소비를 제공함으로써 여러 소비 환경 각각에 QoS를 보장하는 콘텐츠를 서비스할 수 있다.

더욱이 동일한 생산 비용이 여러 개의 전달/소비 환경에 공유될 수 있기 때문에 비용과 효율의 생산성 측면에서도 SVC는 매우 매력적이다. 추후 스케일러블 비디오는 최근의 멀티미디어 서비스 환경에서 보다 향상된 비디오 서비스와 새로운 시장의 기회를 제공할 수 있는 중요한 기술 중 하나일 것이다.

[1] A. Vetro, C. Christopoulos, and H. Sun. "Video transcoding architectures and techniques: An overview", *IEEE Signal Processing Mag.*, vol. 20, pp. 18-29, Mar. 2003.

[2] ITU-T and ISO/IEC JTC1, "Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video," ITU-T Recommendation H.262 - ISO/IEC 13818-2 (MPEG-2), 1994.

[3] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," IEEE Trans. on Circuit System and Video Technology, vol. 11, no. 3, pp. 301-317, Mar. 2001.

[4] ISO/IEC 14496-2, "Information technology-Coding of audio-visual objects - part 2: Visual", International Standard, second edition, December 2001.

[5] MPEG documents, "Registered Responses to the Call for Proposals on Scalable Video Coding," ISO/IEC JTC1/SC29/WG11 MPEG2004/M10569.

[6] ITU-T document, "Joint Draft 9 of SVC Amendment", Joint Video Team JVT-V201, JVT 22nd meeting, Marrakech, Morocco, January, 2007.

[7] ITU-T document, "Joint Scalable Video Model JSVM-9", Joint Video Team JVT-V202, JVT 22nd meeting, Marrakech, Morocco, January, 2007.

[8] ITU-T Recommandation H.264 (2003), "Advanced Video Coding for generic audiovisual services" ISO/IEC 14496-10:2003, "Information technology-Coding of audio-visual objects - part 10: Advanced video coding".

[9] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical B pictures and MCTF," *Proceedings of ICME' 06*, Toronto, Canada, July 2006.

[10] F. Wu, S. Li, and Y.-Q. Zhang, "A framework for efficient progressive fine granular scalable video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 3, pp. 332-344, Mar. 2001.

[11] M. van der Schaar, "A hybrid temporal-SNR fine-granular scalability for internet video," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 3, pp. 318-331, Mar. 2001.

[12] X. Sun, F. Wu, S. Li, W. Gao, and Y.-Q. Zhang, "Macroblock-based progressive fine granular scalable (PFGS) video coding with flexible temporal-SNR scalabilities," Proceedings of ICIP' 01, vol. 2, pp. 1025-

1028, Thessaloniki, Greece, Oct. 2001.

[13] S. Han and B. Girod, "Robust and efficient scalable video coding with leaky prediction," Proceedings of ICIP' 02, Rochester, NY, USA, vol. 2, pp. 41-44, Sep. 2002.

[14] Y. Liu, P. Salama, Z. Li, and E. J. Delp, "An enhancement of leaky prediction layered video coding", IEEE Transaction on Circuits and Systems for Video Technology, vol. 15, no. 11, pp. 1317-1331, Nov. 2005.

[15] Y. Gao and L.P. Chau, "Efficient fine granularity scalability using adaptive leaky factor," IEEE Transactions on Broadcasting, vol. 51, no. 4, pp. 512-519, Dec. 2005.

[16] J.-R. Ohm, "Three-dimensional subband coding with motion compensation," IEEE Transactions on Image Processing, vol. 3, no. 5, pp. 559-571, Sep. 1994.

[17] H. Schwarz, T. Hinz, H. Kirchhoffer, D. Marpe, and T. Wiegand, "Technical description of the HHI proposal for SVC CE1," ISO/IEC JTC 1/SC29/WG11, doc. M11244, Palma de Mallorca, Spain, Oct. 2004.

[18] P. Amon, T. Rathgen, and D. Singer, "File Format for Scalable Video Coding," Joint Video Team JVT-U139, JVT 21st meeting, Hangzhou, China, October, 2006.

[19] M. Wien, R. Cazoulat, A. Graffunder, A. Hutter, and P. Amon, "Real-time System for Adaptive Video Streaming based on SVC," Joint Video Team JVT-U140, JVT 21st meeting, Hangzhou, China, October, 2006.

[20] M. Wien, H. Schwarz, and T. Oelbaum, "Performance Analysis of SVC," Joint Video Team JVT-U141, JVT 21st meeting, Hangzhou, China, October, 2006.

[21] I. Amonou, N. Cammas, S. Kervadec, and S. Pateux, " Optimized Rate-Distortion Extraction with Quality Layer in the H.264/SVC Scalable Video Compression Standard," Joint Video Team JVT-U144, JVT 21st meeting, Hangzhou, China, October, 2006.

[22] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable extension of the H.264/MPEG-4 AVC video coding standard," Joint Video Team JVT-U145, JVT 21st meeting, Hangzhou, China, October, 2006.

[23] Y-K. Wang, M. M. Hannuksela,, S. Pateux, and A. Eleftheriadis, "System and Transport Interface of the Emerging SVC Standard," Joint Video Team JVT-U151, JVT 21st meeting, Hangzhou, China, October, 2006.

[24] S. Wenger, Y-K. Wang, and T. Schierl, "Transport and Signaling of SVC in IP Networks," Joint Video Team JVT-U152, JVT 21st meeting, Hangzhou, China, October, 2006.

최 해 철

1997년 경북대학교 전자공학과(학사)
1999년 KAIST전기 및 전자공학과(석사)
2004년 KAIST전기 및 전자공학과(박사)
2004년 ~ 현재 ETRI 전파방송연구단 방송미디어연구그룹 선임연구원
2007년 ~ 현재 과학기술연합대학원대학교 이동통신 및 디지털방송공학겸임 조교수
관심분야 : 영상통신, 영상분할, 비디오 부호화, Scalable Video Coding, MPEG Video, ITU-T VCEG

강 정 원

1993년 한국항공대학교 항공전자공학과 (학사)
1995년 한국항공대학교 항공전자공학과 (석사)
2003년 Georgia Institute of Technology ECE (박사)
2003년 ~ 현재 ETRI 전파방송연구단 방송미디어연구그룹 선임연구원
2007년 ~ 현재 과학기술연합대학원대학교 이동통신 및 디지털 방송공학겸임 조교수
관심분야 : 비디오신호처리, 비디오 부호화, Scalable Video Coding, 비디오 적응 변환

배 성 준

1997년 고려대학교 전자공학과(학사)
1999년 KAIST전기 및 전자공학과(석사)
2004년 KAIST전기 및 전자공학과(박사)
2004년 ~ 2005년 하나로텔레콤 기술전략팀
2005년 ~ 현재 ETRI 전파방송연구단 방송미디어연구그룹 선임연구원
관심분야 : 비디오 부호화, Scalable Video Coding, MPEG Video, ITU-T VCEG, 비디오스트리밍, 스케일러블스트리밍

유 정 주

1982년 광운대학교 전자통신공학과(학사)
1984년 광운대학교 전자통신공학과(석사)
2001년 영국 LANCASTER 대학교 컴퓨터공학과(박사)
1984년 ~ 현재 ETRI 전파방송연구단 방송미디어연구그룹 방통융합미디어연구팀장
관심분야 : 영상통신, 비디오 부호화, Scalable Video Coding, MPEG Video, ITU-T VCEG, 비디오스트리밍, 스케일러블스트리밍